

WHAT IS CLAIMED IS:

1 1. A method for distributing data among a plurality of data storage
2 systems comprising:
3 obtaining and storing selection criteria;
4 producing profile information for a first data object that is stored in a first data
5 storage system, said profile information comprising content-based information associated
6 with said first data object; and
7 selectively copying said first data object to at least one second data storage
8 system based on said selection criteria and on said profile information,
9 wherein said first data object is copied to said second data storage system
10 depending on content-based information associated with said first data object.

1 2. The method of claim 1 wherein said first data storage system
2 comprises a server component in communication with a data storage component.

1 3. The method of claim 2 wherein said second data storage system
2 comprises a server component in communication with a data storage component.

1 4. The method of claim 1 wherein said selection criteria are stored in said
2 second data storage system, said method further comprising:
3 communicating said profile information to said second data storage system;
4 producing a selection indication based on said selection criteria and on said
5 profile information; and
6 selectively communicating said first data object to said second data storage
7 system based on said selection indication.

1 5. The method of claim 4 wherein said profile information is
2 communicated to a plurality of second data storage systems, said method further comprising:
3 receiving at said first data storage system a selection indication from each of
4 said second data storage systems, wherein said selection indication is an interest metric;
5 producing an ordered set of said second data storage systems, ordered
6 according to said interest metric; and
7 communicating said first data object to the first N of said second data storage
8 systems.

1 6. The method of claim 4 wherein said profile information is
2 communicated to a plurality of second data storage systems, said method further comprising:
3 receiving at said first data storage system a selection indication from each of
4 said second data storage systems, wherein said selection indication is an interest metric;
5 communicating said first data object to a second data storage system if its
6 interest metric exceeds a predetermined threshold.

1 7. The method of claim 4 wherein said profile information is
2 communicated to a plurality of second data storage systems, said method further comprising
3 receiving at said first data storage system a selection indication from each of said second data
4 storage systems, wherein said selection indication indicates whether or not to communicate
5 said first data object to said second data storage system.

1 8. The method of claim 4 wherein if said first data object is not copied to
2 any other data storage system, then determining a replication site from among said other data
3 storage systems independently of content of said first data object and copying said first data
4 object to said replication site.

1 9. The method of claim 1 wherein said selection criteria are stored in said
2 first data storage system, said method further comprising communicating said first data object
3 to said second data storage system based on said profile information and on said selection
4 criteria.

1 10. The method of claim 9 further comprising additional selection criteria
2 for an additional second data storage system, said method further comprising communicating
3 said first data object to said additional second data storage system based on said profile
4 information and said additional selection criteria.

1 11. The method of claim 1 wherein said selection criteria are stored in a
2 selection server system separate from said first data storage system and from said second data
3 storage system, said method further comprising:
4 communicating said profile information to said selection server system;
5 producing in said selection server system a selection indication; and
6 communication said selection indication to said first data storage system,

7 wherein said first data object is selectively communicated to said second data
8 storage system depending on said selection indication.

1 12. A distributed data storage system comprising a plurality of data
2 servers, each data server comprising:

3 a client interface component configured for communication with one or more
4 clients to exchange data;

5 a data storage interface component configured for data communication with
6 data storage component; and

7 a data processing component configured to:

8 produce profile information associated with a first data object that is
9 stored in said data storage component, said profile information comprising content-
10 based information associated with content of said first data object;

11 initiate a comparison of selection criteria with said profile information,
12 said selection criteria comprising criteria associated with at least a second data server,
13 said selection criteria used to determine whether said first data object is copied to said
14 at least a second data server; and

15 copy said first data object to said at least a second data server
16 depending on an outcome of said comparison.

1 13. The data storage system of claim 12 wherein said data processing
2 component is further configured to:

3 communicate said profile information to a plurality of candidate data servers;

4 receive a selection indication from each of said candidate data servers; and

5 copy said first data object to one or more of said candidate data servers based
6 on selection indications received from said candidate data servers,

7 wherein a selection indication is produced by a candidate data server and is
8 based on selection criteria stored in said candidate data server and on said profile
9 information.

1 14. The data storage system of claim 13 wherein said selection indication
2 is a metric that is based on selection criteria and on said profile information.

1 15. The data storage system of claim 13 wherein said selection indication
2 is a binary indicator that indicates whether or not to copy said first data object to said second
3 data server.

1 16. The data storage system of claim 15 wherein said data processing
2 component is further configured to:
3 receive selection criteria from other data servers; and
4 based on said selection criteria and said profile information, selectively copy
5 said first data object to one or more of said other data servers,
6 wherein said other data servers are selected based on selection criteria
7 associated therewith and on said profile information.

1 17. The data storage system of claim 15 wherein said data processing
2 component is further configured to:
3 communicate said profile information to a selection server system that is
4 separate from said data servers;
5 receive selection information from said selection server system; and
6 based on said selection information, copy said first data object to one or more
7 other data servers.

1 18. A method for distributing data among a plurality of data storage
2 systems comprising:
3 obtaining and storing selection criteria in a first data storage system;
4 producing profile information for a first data object that is stored in said first
5 data storage system, said profile information comprising content-based information
6 associated with said first data object; and
7 selectively copying said first data object to at least one second data storage
8 system based on said selection criteria and on said profile information,
9 wherein said first data object is copied to said second data storage system
10 depending on content-based information associated with said first data object.

1 19. The method of claim 18 further comprising receiving, at said first data
2 storage system, said selection criteria from one or more data storage systems other than said
3 first data storage system.

1 20. A data system comprising:
2 a plurality of data centers; and
3 a plurality of client systems in data communication with said data centers,
4 each data center comprising:
5 a data storage component;
6 a file server component operable to exchange data between a client
7 system and said data storage component;
8 a replicator component;
9 a receiver component; and
10 file selection criteria,
11 wherein said replicator component is operable to produce profile data
12 for a data object that is to be replicated among one or more candidate target data
13 centers and to receive a selection indication from each of said candidate target data
14 centers, and to selectively communicate said data object to a candidate target data
15 center based on its selection indication, said profile data representative of content of
16 said data object,
17 wherein said receiver component is operable to receive profile data
18 information from a source data center, said receiver component further operable to
19 communicate a selection indication to said source data center based on said file
20 selection criteria and on said profile data.

1 21. The system of claim 20 wherein said selection indication is an interest
2 metric that is determined based on said file selection criteria and on said profile data, wherein
3 said replicator component is further operable to communicate said data object to a candidate
4 data center based on its interest metric, wherein said candidate target data centers are ordered
5 to produce an ordered set based on their corresponding interest metrics and said replicator
6 component is further operable to communicate said data object to the first N target data
7 centers selected from said ordered set.

1 22. The system of claim 20 wherein said selection indication is an interest
2 metric that is determined based on said file selection criteria and on said profile data, wherein
3 said replicator component is further operable to communicate said data object to a candidate
4 data center based on its interest metric, wherein said replicator component communicates said

5 data object to a candidate target center if its interest metric exceeds a predetermined
6 threshold.

1 23. The system of claim 20 wherein said selection indication is an
2 indication of whether or not to communicate said data object to said candidate target data
3 center.

1 24. A data system comprising:
2 a plurality of data centers; and
3 a plurality of client systems in data communication with said data centers,
4 each data center comprising:
5 a data storage component;
6 a file server component operable to exchange data between a client
7 system and said data storage component;
8 a replicator component; and
9 a collection of selection criteria comprising selection criteria provided
10 from other data centers,
11 wherein said replicator component is operable to produce profile data
12 for a data object that is to be replicated among one or more candidate target data
13 centers and to selectively communicate said data object to said candidate target data
14 centers based on said profile data and selection criteria corresponding to each of said
15 candidate target data centers, said profile data representative of content of said data
16 object.

1 25. The system of claim 24 wherein said replicator module is operable to
2 produce based on said collection selection criteria and on said profile data a plurality of
3 interest metrics, each interest metric corresponding a data center, wherein said candidate
4 target data centers are ordered to produce an ordered set based on their corresponding interest
5 metrics, wherein said replicator component is further operable to communicate said data
6 object to the first N target data centers selected from said ordered set.

1 26. The system of claim 24 wherein said replicator module is operable to
2 produce based on said collection selection criteria and on said profile data a plurality of
3 interest metrics, each interest metric corresponding a data center, wherein said replicator

4 component communicates said data object to a candidate target center if its interest metric
5 exceeds a predetermined threshold.

1 27. A data system comprising:
2 a plurality of data centers, each data center having associated therewith a
3 plurality of client systems; and
4 a selection server system in data communication with said data centers,
5 each data center comprising:
6 a data storage component;
7 a file server component operable to exchange data between a client
8 system and said data storage component; and
9 a replicator component,
10 wherein said replicator component is operable to produce profile data
11 for a data object that is to be replicated among one or more candidate target data
12 centers, to communicate said profile data to said selection server system, and to
13 receive from said selection server system a plurality selection indicators, said profile
14 data representative of content of said data object,
15 wherein said data object is selectively communicated to said candidate
16 target data centers based on said selection indicators,
17 said selection server system comprising a collection of selection criteria
18 comprising selection criteria provided from other data centers, and operable to produce said
19 selection indicators based on said profile data and on said collection of selection criteria.

1 28. The data system of claim 27 wherein said selection server system is a
2 directory server.